

D. Shesternin



Page 1 of 6
#2
Dm
1652

RECEIVED
MAY 03 2001
CENTER 1600/2900

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/466,935

DATE: 04/22/2001
TIME: 13:51:58

Input Set : A:\00101070.txt
Output Set : N:\CRF3\04222001\I466935.raw

ENTERED

```
3 <110> APPLICANT: LIVSHITS, VITALIY
4     ZAKATAEVA, NATALIA
5     ALESHIN, VLADIMIR
6     BELAREVA, ALLA
7     TOKHMAKOVA, IRINA
9 <120> TITLE OF INVENTION: NOVEL GENE AND METHOD FOR PRODUCING L-AMINO ACIDS
11 <130> FILE REFERENCE: 0010-1070-0
13 <140> CURRENT APPLICATION NUMBER: 09/466,935
14 <141> CURRENT FILING DATE: 1999-12-20
16 <150> PRIOR APPLICATION NUMBER: RU98123511
17 <151> PRIOR FILING DATE: 1998-12-23
19 <160> NUMBER OF SEQ ID NOS: 4
21 <170> SOFTWARE: PatentIn version 3.0
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 1231
25 <212> TYPE: DNA
26 <213> ORGANISM: Escherichia coli
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (557)..(1171)
32 <400> SEQUENCE: 1
33 agaaataatg tggagatcgc accgcccac gaatgtgccg gtatatagcg ttacgccac      60
35 ggaccgggct gaacctcctg ctgccagaat gccgccagat catcaacata atcattaaag      120
37 cgattaacat gcccgagatg cggatcggct aacaggcgac cggaacgtcc ctgccgcga      180
39 tggtcgatga ttaagacatc aaaccccaaa tggaacaggt cataggccag ttccgcatat      240
41 ttacgtagc tctcaatacg ccccgggcag atgactacca cccggtcatg gtgctgtgcg      300
43 cgaaaacgga caaagcgcac cggaatgtca tccacaccag taaactctgc ttcatoacgc      360
45 tgacgccaga aatcagtcag cggtccecatg gtaaaagcag caaacgcgtt ttctottggt      420
47 tccagtgctt ttgctgctg aaacatcggg taatctgcct cttaaccac gtaaaatcgt      480
49 ttttttttagc gtgcctgaca caacgtgcg acagtgcgt attgtggcac aaaaatagac      540
51 acaccgggag ttcatac atg acc tta gaa tgg tgg ttt gcc tac ctg ctg aca      592
52           Met Thr Leu Glu Trp Trp Phe Ala Tyr Leu Leu Thr
53           1           5           10
55 tcg atc att tta acg ctg tgc cca ggc tct ggt gca atc aac act atg      640
56 Ser Ile Ile Leu Thr Leu Ser Pro Gly Ser Gly Ala Ile Asn Thr Met
57           15           20           25
59 acc acc tcg ctc aac cac ggt tat ccg gcc ggt ggc gtc tat tgc tgg      688
60 Thr Thr Ser Leu Asn His Gly Tyr Pro Ala Gly Gly Val Tyr Cys Trp
61           30           35           40
63 gct tca gac cgg act ggc gat tca tat tgt gct ggt tgg cgt ggg gtt      736
64 Ala Ser Asp Arg Thr Gly Asp Ser Tyr Cys Ala Gly Trp Arg Gly Val
65 45           50           55           60
67 ggg acg cta ttt tcc cgc tca gtg att gcg ttt gaa gtg ttg aag tgg      784
68 Gly Thr Leu Phe Ser Arg Ser Val Ile Ala Phe Glu Val Leu Lys Trp
69           65           70           75
71 gca ggc gcg got tac ttg att tgg ctg gga atc cag cag tgg cgc gcc      832
72 Ala Gly Ala Ala Tyr Leu Ile Trp Leu Gly Ile Gln Gln Trp Arg Ala
```

RAW SEQUENCE LISTING

DATE: 04/22/2001

PATENT APPLICATION: US/09/466,935

TIME: 13:51:58

Input Set : A:\00101070.txt

Output Set: N:\CRF3\04222001\I466935.raw

```

73          80          85          90
75 gct ggt gca att gac ctt aaa tcg ctg gcc tct act caa tcg cgt cga      880
76 Ala Gly Ala Ile Asp Leu Lys Ser Leu Ala Ser Thr Gln Ser Arg Arg
77          95          100          105
79 cat ttg ttc cag cgc gca gtt ttt gtg aat ctc acc aat ccc aaa agt      928
80 His Leu Phe Gln Arg Ala Val Phe Val Asn Leu Thr Asn Pro Lys Ser
81          110          115          120
83 att gtg ttt ctg gcg gcg cta ttt ccg caa ttc atc atg ccg caa cag      976
84 Ile Val Phe Leu Ala Ala Leu Phe Pro Gln Phe Ile Met Pro Gln Gln
85 125          130          135          140
87 ccg caa ctg atg cag tat atc gtg ctc ggc gtc acc act att gtg gtc      1024
88 Pro Gln Leu Met Gln Tyr Ile Val Leu Gly Val Thr Thr Ile Val Val
89          145          150          155
91 gat att att gtg atg atc ggt tac gcc acc ctt gct caa cgg att gct      1072
92 Asp Ile Ile Val Met Ile Gly Tyr Ala Thr Leu Ala Gln Arg Ile Ala
93          160          165          170
95 cta tgg att aaa gga cca aag cag atg aag gcg ctg aat aag att ttc      1120
96 Leu Trp Ile Lys Gly Pro Lys Gln Met Lys Ala Leu Asn Lys Ile Phe
97          175          180          185
99 ggc tcg ttg ttt atg ctg gtg gga gcg ctg tta gca tcg gcg agg cat      1168
100 Gly Ser Leu Phe Met Leu Val Gly Ala Leu Leu Ala Ser Ala Arg His
101          190          195          200
103 gcg tgaaaaataa tgtcggatgc ggcgtaaacg cttatccga cttactctga      1221
104 Ala
105 205
107 agacgcgtct      1231
110 <210> SEQ ID NO: 2
111 <211> LENGTH: 205
112 <212> TYPE: PRT
113 <213> ORGANISM: Escherichia coli
115 <400> SEQUENCE: 2
117 Met Thr Leu Glu Trp Trp Phe Ala Tyr Leu Leu Thr Ser Ile Ile Leu
118 1          5          10          15
121 Thr Leu Ser Pro Gly Ser Gly Ala Ile Asn Thr Met Thr Thr Ser Leu
122          20          25          30
125 Asn His Gly Tyr Pro Ala Gly Gly Val Tyr Cys Trp Ala Ser Asp Arg
126          35          40          45
129 Thr Gly Asp Ser Tyr Cys Ala Gly Trp Arg Gly Val Gly Thr Leu Phe
130          50          55          60
133 Ser Arg Ser Val Ile Ala Phe Glu Val Leu Lys Trp Ala Gly Ala Ala
134 65          70          75          80
137 Tyr Leu Ile Trp Leu Gly Ile Gln Gln Trp Arg Ala Ala Gly Ala Ile
138          85          90          95
141 Asp Leu Lys Ser Leu Ala Ser Thr Gln Ser Arg Arg His Leu Phe Gln
142          100          105          110
145 Arg Ala Val Phe Val Asn Leu Thr Asn Pro Lys Ser Ile Val Phe Leu
146          115          120          125
149 Ala Ala Leu Phe Pro Gln Phe Ile Met Pro Gln Gln Pro Gln Leu Met
150          130          135          140

```

RAW SEQUENCE LISTING

DATE: 04/22/2001

PATENT APPLICATION: US/09/466,935

TIME: 13:51:58

Input Set : A:\00101070.txt

Output Set: N:\CRF3\04222001\I466935.raw

```

153 Gln Tyr Ile Val Leu Gly Val Thr Thr Ile Val Val Asp Ile Ile Val
154 145 150 155 160
157 Met Ile Gly Tyr Ala Thr Leu Ala Gln Arg Ile Ala Leu Trp Ile Lys
158 165 170 175
161 Gly Pro Lys Gln Met Lys Ala Leu Asn Lys Ile Phe Gly Ser Leu Phe
162 180 185 190
165 Met Leu Val Gly Ala Leu Leu Ala Ser Ala Arg His Ala
166 195 200 205
169 <210> SEQ ID NO: 3
170 <211> LENGTH: 840
171 <212> TYPE: DNA
172 <213> ORGANISM: Escherichia coli
174 <220> FEATURE:
175 <221> NAME/KEY: CDS
176 <222> LOCATION: (187)..(804)
178 <400> SEQUENCE: 3
179 atgccgatca cgcgcagcga aatgctcagc gttaacggcg ttgggatgcg caagctggaa 60
181 cgctttggca aaccgtttat ggcgctgatt cgtgcgcacg ttgatggcga tgacgaagag 120
183 tagtcagcag cataaaaaag tgccagtatg aagactccgt aaacgtttcc cccgcgagtc 180
185 aaatgt atg ttg atg tta ttt ctc acc gtc gcc atg gtg cac att gtg 228
186 Met Leu Met Leu Phe Leu Thr Val Ala Met Val His Ile Val
187 1 5 10
189 gcg ctt atg agc ccc ggt ccc gat ttc ttt ttt gtc tot cag acc gct 276
190 Ala Leu Met Ser Pro Gly Pro Asp Phe Phe Phe Val Ser Gln Thr Ala
191 15 20 25 30
193 gtc agt cgt tcc cgt aaa gaa gcg atg atg ggc gtg ctg ggc att acc 324
194 Val Ser Arg Ser Arg Lys Glu Ala Met Met Gly Val Leu Gly Ile Thr
195 35 40 45
197 tgc ggc gta atg gtt tgg gct ggg att gcg ctg ctt ggc ctg cat ttg 372
198 Cys Gly Val Met Val Trp Ala Gly Ile Ala Leu Leu Gly Leu His Leu
199 50 55 60
201 att atc gaa aaa atg gcc tgg ctg cat acg ctg att atg gtg ggc ggt 420
202 Ile Ile Glu Lys Met Ala Trp Leu His Thr Leu Ile Met Val Gly Gly
203 65 70 75
205 ggc ctg tat ctc tgc tgg atg ggt tac cag atg cta cgt ggt gca ctg 468
206 Gly Leu Tyr Leu Cys Trp Met Gly Tyr Gln Met Leu Arg Gly Ala Leu
207 80 85 90
209 aaa aaa gag gcg gtt tct gca cct gcg cca cag gtc gag ctg gcg aaa 516
210 Lys Lys Glu Ala Val Ser Ala Pro Ala Pro Gln Val Glu Leu Ala Lys
211 95 100 105 110
213 agt ggg cgc agt ttc ctg aaa ggt tta ctg acc aat ctc gct aat ccg 564
214 Ser Gly Arg Ser Phe Leu Lys Gly Leu Leu Thr Asn Leu Ala Asn Pro
215 115 120 125
217 aaa gcg att atc tac ttt ggc tcg gtg ttc tca ttg ttt gtc ggt gat 612
218 Lys Ala Ile Ile Tyr Phe Gly Ser Val Phe Ser Leu Phe Val Gly Asp
219 130 135 140
221 aac gtt ggc act acc gcg cgc tgg ggc att ttt gcg ctg atc att gtc 660
222 Asn Val Gly Thr Thr Ala Arg Trp Gly Ile Phe Ala Leu Ile Ile Val
223 145 150 155

```

RAW SEQUENCE LISTING

DATE: 04/22/2001

PATENT APPLICATION: US/09/466,935

TIME: 13:51:59

Input Set : A:\00101070.txt

Output Set: N:\CRF3\04222001\I466935.raw

```

225 gaa acg ctg gcg tgg ttt acc gtc gtt gcc agc ctg ttt gcc ctg ccg      708
226 Glu Thr Leu Ala Trp Phe Thr Val Val Ala Ser Leu Phe Ala Leu Pro
227      160                      165                      170
229 caa atg cgc cgt ggt tat caa cgt ctg gcg aag tgg att gat ggt ttt      756
230 Gln Met Arg Arg Gly Tyr Gln Arg Leu Ala Lys Trp Ile Asp Gly Phe
231 175                      180                      185                      190
233 gcc ggg gcg tta ttt gcc gga ttt ggc att cat ttg att att tcg cgg      804
234 Ala Gly Ala Leu Phe Ala Gly Phe Gly Ile His Leu Ile Ile Ser Arg
235      195                      200                      205
237 tgatgccaga cgcgtcttca gagtaagtcg gataag      840
240 <210> SEQ ID NO: 4
241 <211> LENGTH: 206
242 <212> TYPE: PRT
243 <213> ORGANISM: Escherichia coli
245 <400> SEQUENCE: 4
247 Met Leu Met Leu Phe Leu Thr Val Ala Met Val His Ile Val Ala Leu
248 1      5      10      15
251 Met Ser Pro Gly Pro Asp Phe Phe Phe Val Ser Gln Thr Ala Val Ser
252      20      25      30
255 Arg Ser Arg Lys Glu Ala Met Met Gly Val Leu Gly Ile Thr Cys Gly
256      35      40      45
259 Val Met Val Trp Ala Gly Ile Ala Leu Leu Gly Leu His Leu Ile Ile
260      50      55      60
263 Glu Lys Met Ala Trp Leu His Thr Leu Ile Met Val Gly Gly Gly Leu
264 65      70      75      80
267 Tyr Leu Cys Trp Met Gly Tyr Gln Met Leu Arg Gly Ala Leu Lys Lys
268      85      90      95
271 Glu Ala Val Ser Ala Pro Ala Pro Gln Val Glu Leu Ala Lys Ser Gly
272      100     105     110
275 Arg Ser Phe Leu Lys Gly Leu Leu Thr Asn Leu Ala Asn Pro Lys Ala
276      115     120     125
279 Ile Ile Tyr Phe Gly Ser Val Phe Ser Leu Phe Val Gly Asp Asn Val
280      130     135     140
283 Gly Thr Thr Ala Arg Trp Gly Ile Phe Ala Leu Ile Ile Val Glu Thr
284 145     150     155     160
287 Leu Ala Trp Phe Thr Val Val Ala Ser Leu Phe Ala Leu Pro Gln Met
288      165     170     175
291 Arg Arg Gly Tyr Gln Arg Leu Ala Lys Trp Ile Asp Gly Phe Ala Gly
292      180     185     190
295 Ala Leu Phe Ala Gly Phe Gly Ile His Leu Ile Ile Ser Arg
296      195     200     205

```

VERIFICATION SUMMARY

DATE: 04/22/2001

PATENT APPLICATION: US/09/466,935

TIME: 13:52:00

Input Set : A:\00101070.txt

Output Set: N:\CRF3\04222001\I466935.raw